Environmental Engineering and Management Journal March 2015, Vol.14, No. 3, 673-680
http://omicron.ch.tuiasi.ro/EEMJ/

"Gheorghe Asachi" Technical University of Iasi, Romania

GRIDDED POPULATION DISTRIBUTION MAP FOR THE HEBEI PROVINCE OF CHINA

Yu Zhang[1,2*], Chun Dong[2], Jiping Liu[2], Shouzhi Xu[3], Tinghua Ai[1], Fengguang Kang[2]

[1]*School of Resource and Environmental Science, Wuhan University, 129 Luoyu Road, Wuhan, 430079 Hubei, China*
[2]*Chinese Academy of Surveying and Mapping, 28 Lianhuachi West Road, Haidian District, 100830 Beijing, China*
[3]*School of Geodesy and Geomatics, Wuhan University, 129 Luoyu Road, Wuhan, 430079 Hubei, China*

Abstract

Mapping the distribution of populations has become an important issue in geographical and relative researchers. Combining population and spatial data allows for socio-graphic information to be visualized, in order to evaluate the total numbers of people at risk of environmental health hazards, who have died in natural disasters etc. Therefore, spatial distribution of population data is an effective way to integrate statistical and spatial data. This paper presents a multi-factor data fusion modeling method for population estimation, which is based on spatial relationships that determine the factors affecting population distribution. The factors that have a strong correlation with population distribution in the Hebei Province were extracted using Geographic Information Systems (GIS). Their standardized weight coefficients were factored as weight coefficients of population distribution in a given spatial unit. The unit (1 km × 1 km) population database was established, allowing for the computation of the relevant population data error. The accuracy of the map was then assessed by comparing predicted population data with that collected from the local government. The results show that the population correlated with geographical factors. The population of the Hebei Province was distributed heterogeneously, increasing from the northwest to southeast. There was relatively low population density in the Taihang Mountains in the west and in the Yanshan Mountains in the northeast, with less than 100 people per square kilometer. The population density in the central Hebei Province was higher, with about 2,000 people per square kilometer, which was higher and denser than that in Handan, Shijiazhuang, Langfang, and Tangshan. These findings may be important for data mining (DM), Decision-making Support Systems (DSS), and regional sustainable development.

Key words: data fusion, grid population, Hebei Province, multi-factor fusion, partial correlation coefficient

Received: September, 2013; Revised final: January, 2015; Accepted: January, 2015

1. Introduction

Extrapolating information from large amounts of data from different sources, such as geographical, statistical, text, and image data is important for policy-making. However, more than eighty percent of all information is related to geographic location in our productive life (Wu, 2009). This information may also be related to different geographical factors in a given spatial unit. Thus, fusing statistical and spatial data provides a scientific basis for administrative planning, land improvement, urban and rural construction, and environmental planning and protection (Balk et al., 2006). Resources, energy, food, environmental issues, and population concerns are among the most pressing global issues today. Resource shortages, environmental degradation, and other problems negatively affect sustainable, social, and economic development (Liu et al., 2006; Small and Cohen, 2004). Apart from playing an increasingly important role in monitoring these changes, Geographic information systems (GIS) also contribute towards regional social development decisions.

Estimations of the total population in irregular areas must be performed quickly and accurately when the need arises, such as in areas affected by nuclear contamination, epidemics, floods, or other

* Author to whom all correspondence should be addressed: e-mail: zhangyu6242@163.com; Phone: +86 10 6388 0559; Fax:+86 10 6388 0540

disasters (Hay et al., 2005; Linard et al., 2010). Several methods have utilized geographical elements and indicators of economic and social development to spatially integrate statistical and quantitative data to address this issue (Liu et al., 2003; Linard et al., 2012). However, these methods are far from ideal. Social statistics primarily focus on the spatial distribution of population-based data (Martin, 2006; Yan and Bian, 2007), while thematic layout charts visualize population and spatial data but are unable to describe population distributions within distinct geographical units or regions.

Current population databases also often fail to take environmental conditions into account. The first (V1) and second (V2) versions of the global population database only use population and administrative boundary data (Zhao et al., 2010), while UNEP/GRID is based only on population density and accessibility (Dobson et al., 2000; Yang et al., 2009). These methods lack precision and are inappropriate for analyzing structured data and environmental conditions, such as topographical features. In northwest China, for instance, the Gobi Desert dominates much of the environment; thus, people tend to settle near oases and the wetlands. Methods using social statistics, however, would fail to identify the significance of the geography in this case. Most existing solutions to this problem focus on selection and quantification of influencing factors in order to take into account the correlation among selected factors. The least squares method has been used to simulate population density in Zhengzhou using population statistics (Lu et al., 2003), while a database of urban layout parameters has also been combined with unit population information for data mining and decision making (He, 2011). These methods include remote-sensing inversion (Jensen et al., 1990), regression-based analysis (Briggs et al., 1997), multi-factor analyses of statistical models (Lloyd, 2010; Niu et al., 1998), Distance Weighted (DW) interpolation, Genetic Programming (GP) and Genetic Algorithms (GA) (Liao et al., 2010), and methods based on night-time imagery and land use data (Zeng et al., 2011), However, few data fusion models can handle the complexity that arises when multiple factors are included (Dong et al., 2000, 2003; Liao, 2005).

The multi-factor data fusion model is the most common approach for analyzing the spatial distribution of social data. In this type of analysis, a multi-factor fusion model is presented based on the weighted coefficients measured between geographical factors and population count to redistribute county unit populations to grid cells in the Hebei Province.

2. Material and methods

2.1. Study area

The area selected for this study was the Hebei Province (longitude: 113°30'-119°54'E, latitude: 36°6'-42°36'N; Fig. 1). The Hebei Province (marked in blue in Fig. 1), an administrative province of China, is located in the Northern China, and embraces capital Beijing and Tianjin municipality. The Hebei Province has an area of 187,693 km^2 and a population of 72.8 million, with 170 administrative counties. The geomorphology is mountainous and hilly with watersheds. The relative elevation difference ranges from about 52 to 2836 meters, with a mean elevation of about 450 meters.

Fig. 1. Location of the study area in China

2.2. Data sources

The data collected from various sources includes population count, a digital elevation model (DEM) and topographical, land cover, and landform data. The population count data for 2010 in the Heibei province were sourced from the 2011 Statistical Yearbook (National Bureau of Statistics of China, 2011). The data was at the county level.

Digital elevation model (DEM) data, topographical features, landform, and land cover datasets were used to reallocate county-based spatial population count data. DEM data were obtained from topographic mapping at the 1:250 000 scale, provided by the Hebei Bureau of Geoinformation, which were capable of reflecting micro-factors in the complex terrain of cities. Topographical features and land cover datasets were obtained from a topographic database at a scale of 1:100 000, provided by National Administration of Surveying, Mapping and Geoinformation (NASMG). Land cover factors included farmland, forests, meadows, water, residential areas, and unused land. Landform data at a scale of 1:250 000 were used by scanning and digitizing the National Physical Atlas of China from SinoMaps Press. Landform factors included plain, hill, platform, mountain, and plateau. The boundaries of administrative divisions, including provincial and county boundaries for Hebei Province at the 1:250

000 scales were obtained from National Basic Geographic Information Center of China.

2.3. Multi-factor data fusion model

2.3.1. Extracting geographical factors

The geographical factors database was established including topographical features, landform, altitude, slope, land cover, and land area by overlaying and analyzing with grid cells. We first created the grid with a 1 km × 1 km cell size and geocoded each cell. Secondly, we extracted the layers, such as inhabitant sites, railways, highways, and waterways from the topographical database, and then overlaid this with the grid created above. Thirdly, the DEM database was used as a data source. According to the relative altitude and slope difference in the Hebei province, classification with the quantile method was taken as a basis to reflect regional differences, and then altitude was reclassified and resampled into nine categories from DEM with the surface analysis function, using ArcGIS10.0 software. Same software was used for obtaining slope.

The spatial analyst function of the software was adopted for resampling and reclassifying slope values into seven categories based on 4° intervals. In the fifth step, the landform layer was overlaid with a grid, and summary statistics for the area of each factor were calculated. Finally, a summary statistics area for each class of land cover was calculated.

2.3.2. Calculating and normalizing geographical factor weighted coefficients

A partial correlation coefficient between the geographical factors and the population was taken as the single-factor weight. The weight coefficient was then normalized for each factor in every county of the Hebei Province in China.

Taking a highway's length factor for example, if there are *n* grids in a county, the normalized length in the i^{th} grid is equal to the total length in the number of *n* grid divided by the highway length in the i^{th} grid. Moreover, the standardized weighted coefficient for highway length in the i^{th} grid is equal to the normalized weight coefficient multiplied by the highway weight coefficient.

The factors such as land cover factors included 6 categories: farmland, forest, meadow, water, residential, and unused land. For these factors, the land cover weight in the i^{th} grid is equal to the sum of the weight coefficient of each category multiplied by the area of each category. Additionally, the standardized weighted coefficient for land cover in the i^{th} grid is equal to the sum of the weight coefficient of land cover in the number of *n* grid divided by the land cover weight in the i^{th} grid. The calculation for the landform factor is the same as that for the land cover factor. The standardized multi-factor weight in the i^{th} grid was obtained from the sum of all factor weights divided by the multi-factor weight in the i^{th} grid, and finally the population in the i^{th} grid is equal to the population count multiplied by the standardized multi-factor weight in the i^{th} grid.

2.3.3. Establishing a data fusion model for population data and geographic factors

Six factors from the geographical databases were selected: topographical factors, altitude, slope, landform, land cover, and land area. These factors were considered important because of their geographical particularity and diversity. Landform factors were omitted to avoid correlations with similar factors, represented as polygons in the databases. The landform factors included in these analyses were those represented in the databases as lines and points (e.g., the entries to habitat sites and the lengths of railways, highways, and waterways).

Geographical factors from the databases were extracted using ArcGIS10.0, in order to generate the data fusion model. The grid module was used to translate factors into a 1 km × 1 km grid. Next, Mask was used to set the Nodata grid to zero for all geographic factors to ensure that results for the geographical factors were equal to actual values. Zonegrid was used to determine regional boundaries, and Valuegrid determined the grids for each geographic factor. Zonal statistics were used to generate an index of geographical factors for each city and county by dividing the final results by 106. This index was used with the Intersect module and overlay function. The average population was calculated by dividing the total population by its total region area. The framework for grid transformation for population data of the Hebei Province is presented in Fig. 2.

2.3.4. Multi-factor data fusion

A multi-factor fusion model was used to examine the relationships between geographical factors and population data. A significant correlation between two variables did not necessarily signify a causal relationship. Two variables can be highly correlated if both are affected by a third shared variable. Consequently, partial correlation coefficients (Fisher, 1924) were used to more accurately portray the relationships between the factors. Partial correlation coefficients (Rao and Sievers, 1995) were used as the weighted coefficients and then normalized (the sum of all weighted coefficients divided by the weighted coefficient). Correlation coefficients generally varied from -1 to 1, and a larger coefficient absolute value reflected a stronger association (Fisher, 1924). Three types of correlation coefficients were used. Pearson correlation coefficients (Dong et al., 2005) were used to analyze continuous variables, such as distance, while Spearman rank and Kendall correlation coefficients were used to analyze ranks.

The population database is divided into 1 km × 1 km grid units for all of the Hebei Province. The weighted coefficient of each grid unit in relation to population distribution W_{mn} is calculated as given by Eq. (1), where: F_{ij} is the area of topographical features for all individual classes, when i=1, it refers to topographical features, and j=1,…, 4 means habitat sites, railways, highways, and waterways; F_{ij} is the altitude area of 9 categories, i=2 means altitude, and j=1, 2…9 means altitude of less than 10m, 10-30 m, 30-50 m, 50-70 m, 70-230 m, 230-350 m, 350-660 m, 660-1100 m, >1100 m; F_{ij} is the area for each all individual slope classes, i=3 means slope, and j=1, 2…7 means slope is in the range 0°-4°, 4°-8°, 8°-12°, 12°-16°, 16°-20°, 20°-24°, 24°-26°; F_{ij} is the landform area of each individual second class, i=4 means landform, and j=1, 2…5 means plain, hill, platform, mountain, plateau; F_{ij} is the land cover area of 6 individual second classes, i=5 means land cover, and j=1, 2…6 means farmland, forest, meadow, water, residential area, unused land; F_{ij} is the land area, i=6 means land cover, and j=1 means land area; b_{ij} is expressed as the standardized weighted coefficient between geographic factors and population, and the range of i and j is equal to F and that of each geographical (sort) factor; a_i is the partial correlation coefficient of the multi-geographical factors with population, and i=1,2…6; a_ib_{ij} is the final weighted coefficient between geographic factors and population; m is the sequence number for each county in the Hebei province; n is the sequence number for each grid in a county.

$$W_{mn} = \frac{a_i b_{ij} F_{ij}}{\sum_{i=1}^{6} \sum_{j=1}^{x} a_i b_{ij} F_{ij}} \tag{1}$$

2.5. Accuracy test

Accuracy tests were conducted by aggregating the gridded population counts to the county level. We then used those counts to produce gridded population distributions and compared the observed population totals at the administrative level with the summed estimates from the output gridded datasets. Root mean square error (RMSE), expressed as a percentage of the mean population size of the administrative level and the mean absolute error (MAE) were used to verify accuracy. This is a better way to compare the estimated population value to the statistical population value and obtain useful information about the quality of our mode (Goovaerts, 2001, 2005; Wang et al., 2005), in order to ensure that the quintiles and probability depend on the interpolation standard errors as much as the predictions. The average standard errors are close to the roots mean squared prediction errors, and the roots mean squared standardized errors should be close to 1. The root mean square error was then calculated with Eq. (2):

$$RMSE=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(P_{si}-P_{ei})^2} \tag{2}$$

where n is the county number in the Hebei province, P_{si} and P_{ei} represent the statistical population value to the estimated population value in i county in the Hebei province.

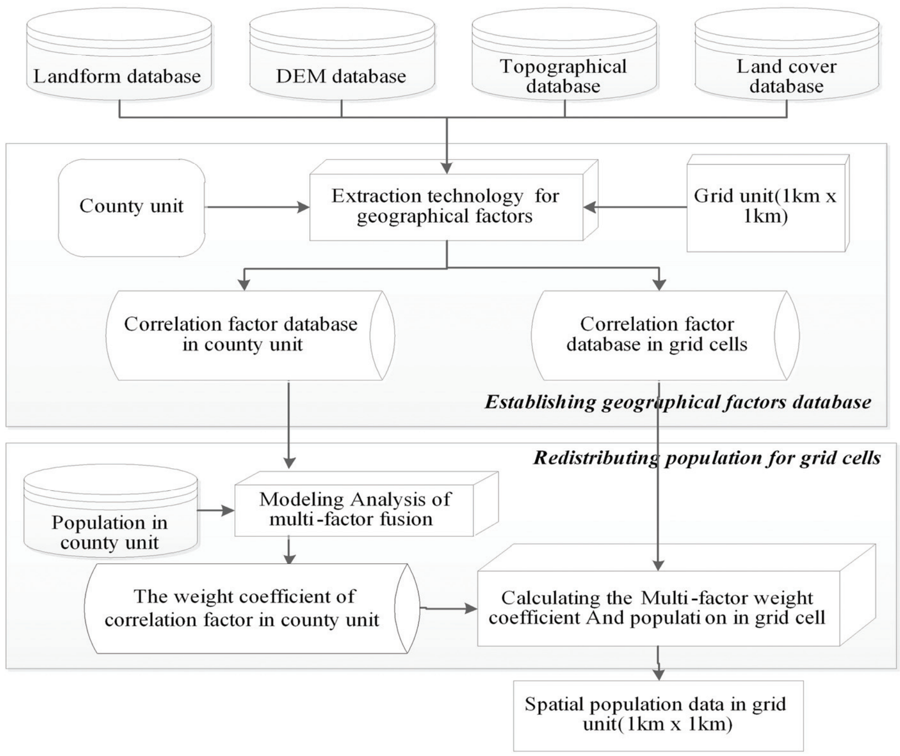


Fig. 2. Framework for establishing a data fusion model for gridded population distribution

3. Results

3.1. Single-factor weight analysis

(1) Topographical features

Inhabited sites, railways, highways, and waterways are the main topographical factors, but are also closely related to the population distribution factor. The quantitative value of a geographical unit is defined as its total length (arc), point size (point), or area (polygon). Polygonal factors, such as residential parcels and rivers, were omitted, and only the parameters represented as lines and points were included in the analysis. Table 1 shows the weighted and standardized weighted coefficients of landforms as they relate to total population data.

(2) Altitude

As altitude increases, population density rapidly declines. In the Hebei Province, population density was higher at lower altitudes (Table 2), and the weighted coefficient was highest for altitudes less than 10 meters. This is consistent with the generalization that population distribution follows a vertical gradient.

(3) Slope

The two rainy areas of the Hebei Province are formed by the slopes of the Yanshan and Taihang Mountains in the north and west, respectively. The population was primarily distributed below a slope of 4° (Table 3). As slope increased, the corresponding weighted coefficient decreased, which is consistent with the principle of upright population distribution.

(4) Landform

The northwest region of the Hebei Province is characterized by mountains, hills, and plateaus, while the southeast and center of the Province are characterized by basins, valleys, and a vast central plain. The population was mainly distributed among the mountains and plains (Table 4).

(5) Land cover

Land cover was highly correlated with the agricultural population of the Hebei Province (Table 5). The population was primarily distributed among farmlands and forests.

(6) Land area

Land area was positively correlated with total population and had a partial correlation coefficient of 0.769.

3.2. Multi-factor weight analysis

In this study, the relative factors were classed with the population distribution of the Hebei Province into six categories: main topographical factors, altitude, slope, landform, land cover, and land area; the weighted coefficients and standardized weighted coefficients (Table 6) were calculated. It is evident that the topographical factors, land cover, and altitude, have the strongest relationship with population in the Hebei Province.

Table 1. The impact of topographical features on population distribution

Topographical feature	*Weighted coefficients*	*Standardized weighted coefficients*
habitat sites (entries)	0.842	0.401
railways (length)	0.674	0.321
highways (length)	0.253	0.120
waterways (length)	0.331	0.158

Table 2. The impact of elevation on population distribution

Altitude (meters)	*Weighted coefficients*	*Standardized weighted coefficients*
<10	0.87	0.207
10-30	0.533	0.127
30-50	0.635	0.151
50-70	0.652	0.155
70-230	0.615	0.146
230-350	0.37	0.088
350-660	0.273	0.065
660-1100	0.143	0.034
>1100	0.119	0.028

Table 3. The impact of slope on population distribution

Slope (°)	*Weighted coefficients*	*Standardized weighted coefficients*
0-4	0.786	0.254
4-8	0.496	0.160
8-12	0.347	0.112
12-16	0.341	0.110
16-20	0.415	0.134
20-24	0.381	0.123
24-26	0.334	0.108

Table 4. The impact of landform on population distribution

Landform (area)	*Weighted coefficients*	*Standardized weighted coefficients*
plain	0.729	0.420
hills	0.199	0.115
platform	0.038	0.022
mountain	0.577	0.333
plateau	0.192	0.111

3.3. Population distribution derived from local governmental offices and calculated using our model

The weighted coefficients between geographical factors and population in the Hebei Province were obtained, and the fundamental geographical factor database covering county administrative boundaries in a 1km x 1km grid were set up. The population database based on a county level was also compiled. Considering the differences in geographical conditions and particularity the distribution of population in different regions of the

Hebei Province, both geographical factor and population databases were helpful for quantitative analysis.

Table 5. The impacts of land cover on population distribution

Land cover (area)	*Weighted coefficients*	*Standardized weighted coefficients*
farmland	0.984	0.243
forest	0.974	0.240
meadow	0.925	0.228
water	0.244	0.060
residential area	0.312	0.077
unused land	0.617	0.152

Table 6. Partial correlation coefficients of the geographical factors

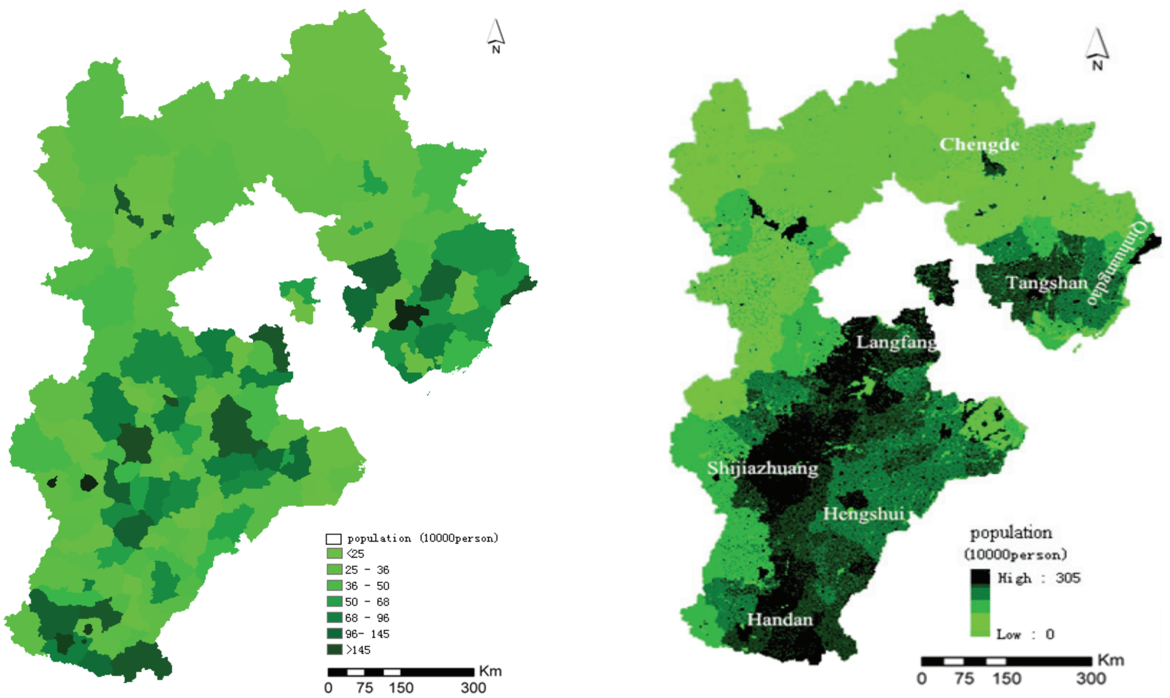
Geographical factors	*Weighted coefficients*	*Standardized weighted coefficients*
topographical feature	0.685	0.158
Altitude	0.638	0.178
Slope	0.753	0.174
Landform	0.771	0.176
Land cover	0.786	0.182
Land area	0.691	0.160

The population data was calculated for both administrative and geographical units using the multi-factor fusion model, which incorporated both spatial and statistical data. The model determined the total population data of distinct regions (Fig. 2a) and compared the results with the population distribution calculated using 1 km grids (Fig. 2b). A darker color indicates a larger population and vice versa. The average population distribution within each region did not accurately characterize that of the same region when population distribution was calculated by grid units. The population clustered near administrative centers, such as in Zhangjiakou, Chengde, Langfang, Qinhuangdao, and Hengshui, which show a higher population concentration of about 2000 persons per square kilometer. These regional trends were not reflected by the regional population distribution, however. The population distribution gradually increased from the northwest to southeast, most likely due to a higher slope in the northwest. The northwest also has mountains, hills, and plateaus, while the central and southeast regions are comprised of basins, valleys, and plains. Population density was relatively low in the Taihang Mountains and was less than 100 persons per square kilometers for the Yanshan Mountains.

The population density in the central Hebei Province was higher, with about 2,000 people per square kilometer. Overall, the population distribution in Hebei is higher and denser than that seen in Handan, Shijiazhuang, Langfang, and Tangshan. In addition, the spatial population distribution was uneven.

3.4. Accuracy assessment

Mapping accuracies are consistently higher when incorporating geographical factor information. Since spatially detailed county data for the Hebei Province were available, we aggregated the small administrative units into a coarser administrative unit level by summing the small units.



a) b)

Fig. 2. Spatial population distribution maps in the Hebei Province, China (a) Population derived by municipality; b) Population calculated with our model

Then, these coarse units and population sums were used to generate gridded population maps and the sums of those gridded estimates were compared with the numbers from the original unit population (Fig. 3). The population statistical error is in the range of 6.84 percent. The average standard errors are close to the root mean squared prediction errors, and the root mean squared standardized error was 0.94, which is quite close to 1. Compared with calculating the average population in all counties, the population accuracy with spatial population distribution in grid cells improved to a certain extent; the overall average error was less than the simulation based on land use (Tang et al., 2012; Tian et al., 2004).

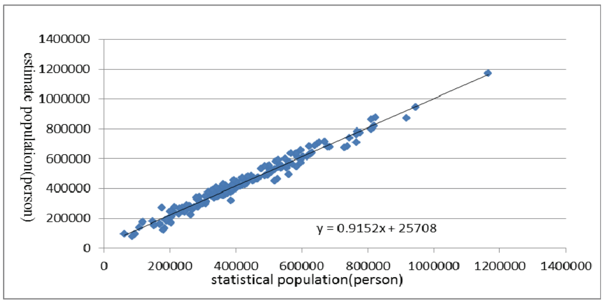


Fig. 3. Scatter diagram of estimated population value and statistical population value

4. Conclusions

In order to analyze the relationship between geographical factors and population data in a spatial unit, in this study, the geographical factors database was established with GIS software, including topographical features, landform, altitude, slope, land cover, and land area. Based on the analysis of geographical factors and population by partial correlation coefficient, there was a significant correlation between population and geographical factors. However, we must be noted that the effect of each factor on population is not the same as that in other regions due to China's topographical features.

The gridded dataset presented here more accurately characterizes population distribution in the defined regions than other existing global datasets (e.g., Gridded Population of the World) with small-scale data. The gridded population dataset takes advantage of the growing collection of geographical features and land cover data to more accurately map human population distributions at a finer spatial scale. Additionally, weighting the distribution of a population by different land cover types, especially through incorporating detailed datasets on roads and developed areas, provides a more accurate representation of population density.

The results prove that it is much more effective than other methods in counting the population in an arbitrary region in China. Measuring population distribution is very complicated, since it relies on the distribution of other statistical factors, some of which are hard to quantitate. There are multiple other factors affecting population distribution, as well.

Analyzing statistical data in an irregular and random geographical region is important for policy making. Thus, spatializing population data can help develop new ideas and models for space-time data integration and the fusion of multi-source and multi-scale data. In addition, social-pixel and pixel-society concepts can be cemented, promoting a social and public service-oriented e-government. These problems affect the accuracy of the calculated population distribution and should be improved in our future studies.

Acknowledgements

This research was funded by the National Key Technology R&D Program under grant No.2012BAH28B03, and by National High Technology Research and Development Program of China (863 Program) under grant No. 2012AA12A402. We wish to thank the invited reviewers for their time and comments, which improved the manuscript.

References

Balk D.L., Deichmann U., Yetman G., Pozzi F., Hay S.I., Nelson A., (2006), Determining global population distribution: methods, applications and data, *Advances in Parasitology*, **62**, 119-156.

Briggs D.J., Collins S., Elliott P., Fischer P., Kingham S., Lebret E., Pryl K., Reeuwijk H.V., Smallbone K., Veen A.V.D., (1997), Mapping urban air pollution using GIS a regression-based approach, *International Journal of Geographical Information Science*, **11**, 699-718.

Dong C., Zhang Q.P., Zhang J.Q., Liu J.P., Wang L., (2000), The application research of establishing geographical parameter (in Chinese), *Remote Sensing Information*, **1**, 12-17.

Dong C., Zhao R., Liu J.P., (2003), An application of geographical parameters database in establishment of unit population database, *Chinese Geographical Science*, **13**, 34-38.

Dong C., Luo Y.B., Liu J.P., Wu X.Z., Wang G.X., (2005), Study on correlation between residential points and geographical factors based on poisson logarithm linearity model (in Chinese), *China Population, Resources and Environment*, **15**, 79-84.

Fisher R.A., (1924), The Distribution of the Partial Correlation Coefficient, *Metron*, **3**, 329-332.

Goovaerts P., (2001), Geostatistical assessment and validation of uncertainty for three-dimensional dioxin data from sediments in an estuarine river, *Environmental Science and Technology*, **35**, 3294-3301.

Goovaerts P., (2005), Geostatistical modeling of the spatial variability of arsenicin groundwater of southeast Michigan, *Water Resources Research*, **41**, 1-19.

Hay S.I., Noor A.M., Nelson A., Tatem A.J., (2005), The accuracy of human population maps for public health application, *Tropical Medicine and International Health*, **10**, 1073-1086.

He L.N., (2011), A discussion of population spatial distribution based on urban layout parameter (in

Chinese), *Science of Surveying and Mapping*, **36**, 38-41.

Dobson J.E., Bright E.A., Coleman P.R., Durfee R.C., Worley B.A., (2000), LandScan: a global population database for estimating populations at risk, *Photogrammetric Engineering & Remote Sensing*, **66**, 849-857.

Jensen J.R., Ramsey III E.W., Holmes J.M., Michel J.E., Savitsky B., Davis B.A., (1990), Environmental sensitivity index (ESI) mapping for oil spills using remote sensing and geographic information system technology, *International Journal of Geographical Information Systems*, **4**, 181-201.

Liao Y.L., (2005), *The study for-scaling with socio-economic data,* (in Chinese), MSc Thesis, Nanjing Normal University, Nanjing, China.

Liao Y.L., Wang J.F., Meng B., Li X.H., (2010), Integration of GP and GA for mapping population distribution, *International Journal of Geographical Information Science*, **24**, 47-67.

Linard C., Alegana V.A., Noor A.M., Snow R.W., Tatem A.J., (2010), A high resolution spatial population database of Somalia for disease risk mapping, *International Journal of Health Geographics*, **9**, 1–13.

Linard C., Gilbert M., Snow R.W., Noor A.M., Tatem A.J., (2012), Population distribution, settlement patterns and accessibility across Africa in 2010, *PLoSONE*, **7**, 1-8.

Liu J.P., Liu Z., Wang L., (2006), Research on the spatial information service for e-government based on function collaboration (in Chinese), *Acta Geodaetica et Cartographica*, **35**, 299-302.

Liu J.Y., Yue T.X., Wang Y.A., Qiu D.S., Liu X.Z., Deng M.L., Yang X.H., Huang Y.J., (2003), Digital simulation of population density in China, (in Chinese), *Acta Geographica Sinica*, **58**, 17-24.

Lloyd C.D., (2010), Exploring population spatial concentrations in Northern Ireland by community background and other characteristics: an application of geographically weighted spatial statistics, *International Journal of Geographical Information Science*, **24**, 1193-1221.

Lu A.M., Li C.M., Lin Z.J., Shi W.Z., (2003), Spatial continuous surface model of population density, (in Chinese), *Acta Geodaetica et Cartographica*, **32**, 344-348.

Martin D., (2006), *Grid models of population: temporal comparison by fixing the geography*, ESRC Research Methods Festival, University of Southampton, On line at: http://slideplayer.com/slide/786548/.

National Bureau of Statistics of China, (2011), *Statistical Yearbook of Hebei* (in Chinese), China Statistics Press, Beijing, China.

Niu H.E., Meng Q.M., Hu Q.C., Chen Y.C., (1998), Economic interaction analysis between regions of Gansu Province and their surrounding areas (in Chinese), *Economic Geography*, **18**, 51-56.

Small C., Cohen J.E., (2004), Continental physiography, climate, and the global distribution of human population, *Current Anthropology*, **45**, 269-289.

Rao S.H., Sievers G.L., (1995), A robust partial correlation measure, *Journal of Nonparametric Statistics*, **41**, 1-20.

Tang Q., Xu X.Y., Yu S., Xin D., (2012), Population spatial distribution and its application based on GIS: Case of northern China, *Journal of Beijing Normal University (Natural Science)*, **48**, 654-659.

Wang C.J., Tang X.H., (2004), GIS-based Specialization of population census data in Fujian Province (in Chinese), *Geography and Geo-Information Science*, **20**, 71-74.

Wang H., Liu G., Gong P., (2005), Use of co-kriging to improve estimates of soil salt solute spatial distribution in the yellow river delta, *Acta Geographica Sinica*, **60**, 511-518.

Wu X.C., (2009), *Principal and Method of Geographical Information Systems* (Second Edition), Publishing House of Electronics Industry, Beijing, China.

Yan Q.W., Bian Z.F., (2007), Method of pixelizing social statistical data based on the GIS, (in Chinese), *Yunnan Geographic Environment Research*, **19**, 92-97.

Yang X., Huang Y., Dong P., Jiang D., Liu H., (2009), An updating system for the gridded population database of china based on remote sensing, GIS and Spatial Database Technologies, *Sensors*, **9**, 1128-1140.

Zeng C.Q., Zhou Y., Wang S.X., Yan F.L., Zhao Q., (2011), Population spatialization in China based on night-time imagery and land use data, *International Journal of Remote Sensing*, **32**, 9599-9620.

Zhao J., Yang D. H., Pan J.H., (2010), A study on spatial pattern of GDP in Lanzhou city based on spatialization and land utilization, (in Chinese), *Journal of Northwest Normal University (Natural Science)*, **46**, 92-97.